Savona, Maria

Article

Data governance: Main challenges

EconPol Forum

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Savona, Maria (2024) : Data governance: Main challenges, EconPol Forum, ISSN 2752-1184, CESifo GmbH, Munich, Vol. 25, Iss. 3, pp. 28-31

This Version is available at:
https://hdl.handle.net/10419/298821

Maria Savona

Data Governance: Main Challenges*

KEY MESSAGES

- **Addressing the governance of emerging digital automation technologies and data in particular requires a multidisciplinary perspective, including techno-legal, geopolitical and economic expertise**
- **Research on governing the process of individual and B2B data sharing, either through mandatory rules or the creation of incentives for sharing, will be important for setting the policy agenda**
- **A "data-haven hypothesis" might explain asymmetries in the concentration of digital infrastructure, with countries with more stringent data protection, IP or tax regimes offshoring cloud services and data hubs to countries with weaker ones**
- **The EU AI Act might lead to a new wave of the so-called "Brussels effect," even though it may still not be optimal and require further debate and public scrutiny**

Economists of innovation know too well that the governance of emerging technologies to prevent potential side effects of uncontrolled developments usually requires more time than firms need to enter those markets. The unprecedented pace of development of digital automation technologies and artificial intelligence (AI) makes the identification of such effects and the formulation of tools to address these challenges complex from different perspectives.

The first one is *techno-legal* and concerns the pervasiveness of AI applications and the need to regulate them in very diverse realms, often at odds with each other (e. g., the attribution of intellectual property rights on AI-generated art; the protection of privacy in increasingly complex data-treating business models). The second one is *geopolitical* and specific to AI, which seems to have sparked a wave of "new protectionism" and ensuing tensions among China, the US and the EU, on pretty much every aspect related to digitalization, from domestic chipmaking to the regulation of digital trade and cross-border data flows "with trust" (OECD 2022). The third one is *economic* and includes, for instance, the need to adapt and possibly "upgrade" competition and antitrust regulations to digital markets; mitigate the effects of digital automation on labor markets; ensure a fair and inclusive redistribution of both the private and social value generated by (personal and business) data among firms, individual data subjects and public actors.

The case of generative AI is an example of the extent to which we shall understand and predict how the emerging digital automation technologies raise questions that have been unprecedented in the history of other technological paradigms. Never have the same entrepreneurs and innovators, owners of "too-big-to-fail" platforms, demanded regulatory intervention from governments to "slow down" the development of generative AI,[1] the core of their business and competitive advantage. Neither have they explicitly expected public institutions to identify and regulate undesirable effects such as fake news and cybersecurity.

Addressing each of the above challenges and understanding how they are interrelated is an arduous task. We offer a brief reflection on two – relatively less explored – policy-relevant economic aspects of data governance, data sharing and the concentration of digital infrastructure, and then focus briefly on the recent EU AI Act.

Maria Savona

is Professor of Economics of Innovation at the Science Policy Research Unit (SPRU) at the University of Sussex, UK and Professor of Economics at the Department of Economics and Finance at LUISS University, Rome. She coordinates the Industrial Policy Area within LEAP (Luiss Institute for Economic Analysis and Policy). She is also the Coordinating Editor for Research Policy.

DATA SHARING

The economic nature of data changes along the data "value chain," which includes the aggregation, processing and analytics of individual data (Corrado et al. 2022; Goos and Savona 2024). Individual data[2] is a *club good*, excludable but not rivalrous (Savona 2019), as individuals or business might prevent the use of their personal or copyright-protected[3] information. However, once shared, data can be re-used at virtually no marginal costs. A legally owned database is a *private good*, excludable, and rivalrous, and is usually included in the intangible assets of firms (Corrado et

* The paper builds on several solo and joint working and briefing papers, keynote addresses and panel discussions over the past few years, quoted in the text and referenced.

[1] See "Pause Giant AI Experiments: An Open Letter" (March 2023): https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

[2] Personal data means "any information relating to an identified or identifiable natural person ("data subject"); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" (article 4(1), EU GDPR, 2018).

[3] EU Directive 96/9/EC of 11 March 1996 recognizes the legal ownership of databases to firms, with *database property rights* being a legal category implemented in that context.

al. 2022), being thus a source of comparative advantage. The ensuing data analytics is valuable information that eventually becomes collective knowledge whose economic nature is inherently a *public good*.

Depending on the actors involved and the purpose that information and collective knowledge serve, data presents the challenge of having to reconcile objectives that are often at odds with each other. For instance, it is important to create incentives to *maximize data sharing* for purposes of public interest such as health, mobility, or research. However, data as an asset in firms that benefit from inherent network economies require *capping private value concentration* from an antitrust perspective. Facilitating data sharing and preventing value concentration might be at odds with *protecting individual privacy and other rights* (Savona 2020 and 2021; Goos and Savona 2024). The European Commission has been trying to resolve this policy conundrum in the context of the articulated regulatory framework developed over the past few years and considered a benchmark worldwide.

An interesting instance of such EU regulations is the EU Data Governance Act (DGA), which has explicitly aimed to foster the "availability of data for use by *increasing trust in data intermediaries* and *by strengthening data sharing mechanisms* across the EU." The focus is on the *creation of data markets* by legitimizing data intermediaries (i. e. , data trusts, cooperatives, stewards, unions). Further, it aims to *"make public sector data available for re-use (..) on altruistic grounds."*

Data intermediaries are supposed to act in the interests of individual data subjects and facilitate data sharing (Savona 2021; Goos and Savona 2024). However, to achieve a sufficient scale of aggregate information that serves public purposes such as research and public health, data intermediaries would need large-scale digital infrastructure to manage large amounts of data, which might lead to the same challenges that current big techs pose, such as market concentration, privacy leakages, and cybersecurity.

In addition, trustees that operate on a fiduciary basis on behalf of a group of individual data subjects should demonstrate a commitment to pro-social and "altruistic" behavior, supported by appropriate incentives. This is not trivial.

A governance model that enforces data sharing for public interest has been proposed for the design and launch of the green mobility plan of the City State of Hamburg (The New Institute 2023). Within the legal framework designed in this case, data sharing has been made mandatory, rather than delegated to voluntary data trusts. The effectiveness of the DGA in creating missing data markets through data intermediaries is yet to be assessed, but it would be important that the intermediaries be capped in scale, limited to specific purposes, and monitored by an independent governing body in order to minimize risks of shifting from big tech to big trusts.

Graef and Prufer (2021) propose a governance framework for B2B data sharing that aims at avoiding market concentration. From a legal perspective, they claim that data sharing should be made mandatory and regulated, and propose three potential models.

The first model would be a fully centralized one, involving a central role for a European Data Sharing Agency that would manage a mandatory data sharing. The second model would be fully decentralized, involving the creation of a Data Sharing Cooperation Board, which would oversee a network of National Competition Authorities (NCAs) whose remit would be to enforce data sharing contracts. The third one would be a hybrid model, with both centralized and decentralized features.

Governing the process of individual and B2B data sharing, either through mandatory rules or the creation and maintenance of incentives for sharing that do not lower consumer and citizens' protection, is no easy task. Overall, research and case studies on the creation and implementation of regulatory frameworks with different degrees of centralization are still in their infancy, let alone the assessment of their effectiveness. This is likely to become an intriguing research and policy agenda in the near future.

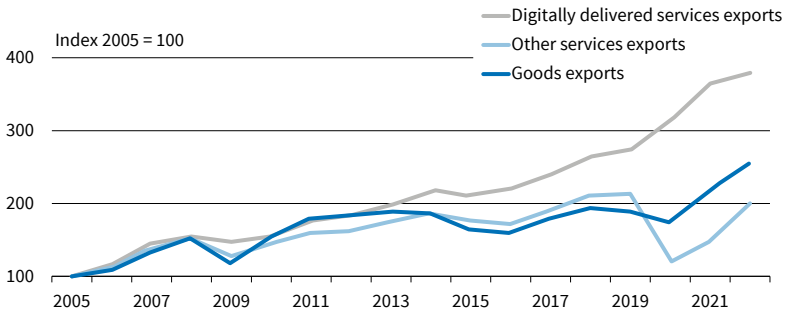
THE GEOPOLITICS OF DIGITAL INFRASTRUCTURE

Trade in digital services has increased considerably over the past decades (Figure 1), and relies on the investment capacity in physical digital infrastructure that supports cross-border data flows, including submarine cables, optic fibers, and, more recently, data centers and cloud storage of data and software. Arguably, the intertwined effect of technological advances in digitalization and the specificities of the digital infrastructure needed to support cross-border data flows are changing the sources of comparative advantage of countries in the digital service trade.

According to IMF et al. (2023), *"cloud computing services, defined as 'computing, data storage, software, and related IT services accessed remotely over a network, supplied on demand and with measured*

Figure 1

Growth of Goods, Services and Digitally Delivered Services Exports



Note: Digitally delivered services include GATS mode 1 exports of financial, insurance, telecommunications, computer and information services (ICT), charges for the use of intellectual property, and most of other business services and of personal, cultural and recreational services in the Balance of Payments.
Source: Trade Finance Global (2023); Papadakis and Savona (2024).

resource usage that allows charging on a pay-per-use basis', are increasingly used to replace ownership of on-premises IT equipment." This means that, particularly when the scale of digital activity increases, the costs of storing and processing data lead companies to outsource (and offshore) data stocks to external cloud service providers and data centers.

Papadakis and Savona (2024) show that digital infrastructure (data centers and cloud storage) is unevenly concentrated across developed and developing countries, with a non-negligible share located in small developing countries. There are different potential explanations for this.

First, the concentration of digital infrastructure might mirror the asymmetrical distribution of (digital) trade among headquarter and factory countries (Baldwin and López-González 2015), with large core countries offshoring digital infrastructure to peripheral and small economies, reproducing a core-periphery structure of digital trade.

Second, a high concentration of digital infrastructure in specific countries might be due to different digital regulatory regimes, including the articulated EU digital regulations mentioned in the previous section, the EU adequacy regulations on digital trade (see e. g. , Ferracane et al. 2023; Bacchus et al. 2024), and intellectual property (IP) regulatory regimes (Santancreu 2023). Data storage might be concentrated in countries that are destination of IP profit shifting or patent boxes (Haufler and Schindler 2023; Alstadsæter et al. 2018; Accoto et al. 2023).

In Papadakis and Savona (2024) we argue that a "data-haven hypothesis" might explain asymmetries in the concentration of digital infrastructure, similarly to how the "pollution-haven hypothesis" has explained patterns of trade of green and brown products: advanced countries offshore activities that would not meet their strict environmental regulations to mid- and low-income countries with less stringent regulations (see Savona and Ciarli 2020 for a selected review). In the same vein, countries with more stringent data protection, IP or tax regimes would offshore cloud services and data hubs to countries with weaker ones.

The idea of increasing "data governance interoperability" (Bacchus et al. 2024) might go in the direction of strengthening the role of national governments vis-à-vis private owners of data centers or cloud services. However, the plea for international cooperation to ensure interoperability of data governance regimes should be extended beyond data protection to other realms, including IP and tax regulation.

THE EU AI ACT

The European regulatory framework of digital emerging technologies has always been at the forefront of what has been named the "Brussels effect." When the GDPR became law, US tech giants had to comply, and several governments chose to align themselves onto the main principles and rules to protect citizens' privacy – and digital rights – more broadly. It will be interesting to see whether the EU Artificial Intelligence Act will trigger another Brussels effect. A few considerations are in order.

First, since the GDPR, the development of AI applications, the market concentration and the lobbying of US Big-Tech now calls for articulated and comprehensive governance of data and AI that goes well beyond individual privacy protection. As mentioned above, governance interoperability (Bacchus et al. 2024) can be fostered by reducing the widening gaps in digital, IP and tax regulations.

The EU AI Act includes not only a systematization of high-risk cases, such as predictive policing, social scoring, and algorithmic management in workplaces, but also an attempt to regulate foundation models such as LLMs, which have sparked much debate in the case of generative AI. As it has been pointed out, the regulation of foundation models is at the root of AI governance, and this is essentially what will be at stake over the next few years.

This opens a Pandora box and leads to a second point: there seem to be hints that the US is moving closer to the EU's regulatory framework. One of the issues at stake is the alleged copyright infringement on digital texts copied from the web and used to train LLMs and generative AI. It is well known how the debate has been nurtured by the cases of the *New York Times* and, separately, eight other American newspapers owned by Alden Global Capital – including the *Chicago Tribune* and *New York Daily News* – suing OpenAI and Microsoft. In the *New York Times* instance, the complaint crucially goes beyond the infringement of copyright law and lays down the case for regulating AI more broadly, borrowing much of the thrust and the principles of risk-adverse and rights-preservation contained in the EU AI Act. It raises concerns that touch upon misinformation, the protection of human creativity, the social value of professional and truthful journalism, as well as democracy itself. A highly reputable US company is suing a formerly non-profit and now for-profit billion-heavy US company.

A further instance where the US has moved quite unexpectedly toward the EU regulatory framework is in the sudden change of its position on digital trade (Ruiz and Savona 2024). The US announced last October that it was withdrawing its position on digital trade from the WTO to allow for stronger regulation. This might certainly be in line with the protectionism strategy in the context of geopolitical tensions mentioned above and the wish to maintain the US forefront position in the global AI race. However, it is not inconsistent with the Biden administration's Blueprint for an AI Bill of Rights.

In sum, the EU AI Act might still not be optimal and may require further debate and public scrutiny.

However, it may still lead to a new wave of the Brussels effect, as the governance of AI and data is and will continue to be increasingly challenging.

CONCLUDING REMARKS

As briefly argued above, one of the challenges of AI and data governance is to reconcile often conflicting objectives: to create (and maintain) incentives to maximize data sharing for purposes of public interest, such as health or research; to limit the concentration of private value arising from (involuntary or voluntary) data collection and analytics as in the case of LLM training; to protect privacy and other individual rights such as copyright in a context where human creativity (still) has social value.

All this calls for thinking out of the box, relying on a multidisciplinary understanding of: (i) what the (economic) detrimental effects of a badly or non-regulated technology are, linked with (ii) carefully designed legal frameworks that prevent or internalize these externalities, alongside a (iii) forward-looking view of how the geopolitics of technology and the striking asymmetries in the lobbying powers of different actors involved play out.

REFERENCES

Accoto, N., S: Federico and G. Oddo (2023), "Trade in Services Related to Intangibles and the Profit Shifting Hypothesis", *Temi di discussione (Economic Working Papers)* 1414, Bank of Italy, Economic Research and International Relations Area.

Alstadsæter, A., S. Barrios, G. Nicodeme, A. M. Skonieczna and A. Vezzani (2018), "Patent Boxes Design, Patents Location, and Local R&D", *Economic Policy* 33(93), 131–177.

Bacchus, J., I. Borchert, M. Morita-Jaeger and J. Ruiz Macpherson (2024), *Interoperability of Data Governance Regimes: Challenges for Digital Trade Policy,* CITP Briefing Paper 12, https://citp.ac.uk/publications/interoperability-of-data-governance-regimes-challenges-for-digital-trade-policy.

Baldwin, R. and J. López-González (2015), "Supply-chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses", *The World Economy* 38, 1682–1721.

Corrado, C., J. Haskel, M. Iommi, C. Jona-Lasinio and F. Bontadini (2022), *Data, Intangible Capital and Productivity*, https://www.nber.org/books-and-chapters/technology-productivity-and-economic-growth/data-intangible-capital-and-productivity.

European Commission (2022), *European Data Governance Act*, https://digital-strategy.ec.europa.eu/en/policies/data-governance-act.

European Commission (2023), *Data Act*, https://digital-strategy.ec.europa.eu/en/policies/data-act.

Ferracane, M., B. Hoekman, E. van der Marel and F. Santi (2023), "Digital Trade, Data Protection and EU Adequacy Decisions", *CIP Working Paper* 6.

Goos, M. and M. Savona (2024), "The Governance of Artificial Intelligence: Harnessing Opportunities and Mitigating Challenges", *Research Policy* 53(3), 104928.

Graef, I. and J. Prüfer (2021), "Governance of Data Sharing: A Law & Economics Proposal", *Research Policy* 50(9), https://doi.org/10.1016/j.respol.2021.104330.

Haufler, A. and D. Schindler (2023), "Attracting Profit Shifting or Fostering Innovation? On Patent Boxes and R&D Subsidies", *European Economic Review* 155, 104446, https://doi.org/10.1016/j.euroecorev.2023.104446.

IMF, OECD, UN and WTO (2023), *Handbook on Measuring Digital Trade*, 2nd Edition, https://www.wto.org/english/res_e/publications_e/digital_trade_2023_e.htm.

Papadakis, I. and M. Savona (2024), *The Uneven Geography of Digital Infrastructure. A 'Data Haven Hypothesis'?*, CITP Briefing Paper, Forthcoming.

Ruiz, J. and M. Savona (2023), "The US Turn Is Reshaping the Geopolitics of Digital Trade. What Does This Mean for the UK?", *CITP Blog*, December.

Santacreu A. M. (2023), "International Technology Licensing, Intellectual Property Rights and Tax Havens", *The Review of Economics and Statistics* 2023, https://doi.org/10.1162/rest_a_01382.

Savona, M. and T. Ciarli (2019), "Structural Changes and Sustainability. A Selected Review of the Empirical Evidence", *Ecological Economics* 159(c), 244–260.

Savona, M. (2019), "The Value of Data: Towards a Framework to Redistribute It", *SPRU Working Paper Series (SWPS)* 2019-21, www.sussex.ac.uk/spru/swps2019-21.

Savona, M. (2020), "Governance Models for Redistribution of Data Value", *VOX CEPR* 17 January.

The New Institute New Hanse Blueprint (2023), *Governing Urban Data for the Public Interest*, Hamburg.

Trade Finance Global (2023), *Global Trade Shows Resilience (for Now) in the Face of Increased Fragmentation: Standard Chartered Global Research Special Report*, https://www.tradefinanceglobal.com/posts/global-trade-shows-resilience-now-face-of-increased-fragmentation-standard-chartered-global-research-special-report/.